图书情报工作
第 63 卷 第 7 期 2019 年 4 月

知识组织生态系统构架形成与研究进展*

常春[1] 杨婧[2] 李永泽[3]
[1] 中国科学技术信息研究所 北京 100038 [2] 中国版本图书馆 北京 100005
[3] 南京大学信息管理学院 南京 210023

摘要：[**目的/意义**]将生态学的原理与方法引进到知识组织系统研究中，以更好地发挥知识组织系统的功能和作用。[**方法/过程**]基于生态学的研究成果，将生态学的原理和方法引入到知识组织系统的构建和应用中，生物个体与概念实例对应，物种种群与概念对应，生物群落与词表对应，生物环境与文献环境对应，在结构和功能完整一致对应基础上，研究知识组织系统的工作原理和方法。[**结果/结论**]分概念实例、概念种群、概念群落(词表)和生态系统 4 个层面，细化和完善知识组织生态系统构架，分析总结各个层面的主要研究内容。

关键词： 知识组织生态系统 生态学 知识组织系统 叙词表
分类号： G254
DOI：10.13266/j.issn.0252 – 3116.2019.07.017

1 引言

生态学(ecology)是研究有机体及其周围环境相互关系的科学[1]。2015 年生态学的原理和方法被引入到知识组织体系构建和应用中，常春[2]提出知识组织生态系统概念(KOES, knowledge organization ecosystem)，经过其研究团队将近 3 年的研究和实践，通过概念生态位的研究[3]，基于能量流动的概念等级关系研究[4]，基于种间关系的概念相关关系分类研究等[5]，对其构架又进行了细化，在组成方面，由以前的物种加环境构成生态系统，细化为生物个体、生物种群、生物群落，外加生物环境基础上构成生态系统；对应到知识组织系统(KOS, knowledge organization system，本文也简称为“词表”)，同样细化为实例、概念、词表，外加文献环境基础上构成知识系统。将生态系统的原理与方法应用到知识系统中，在结构和功能上形成了知识组织生态系统。研究与应用内容也从以前的 3 个组面，细化为现在的个体、概念、词表、知识系统 4 个层面，与生态学的学科构架达到完全的对应统一，本文主要介绍具体细化内容与研究进展。

2 知识组织生态系统总体构架

2.1 知识组织生态系统结构与组成

经过细化的知识组织生态系统，其形成过程见图 1。同一物种不同生物个体(individual)组成种群(population)，种群是生物物种的基本单位；不同的物种组成群落(community)，群落是植物、动物、微生物等全体生物的总称；由生物群落与其全部生存环境组成生态系统(ecosystem)。与生态学相对应，在知识组织系统中，由不同的实例组成概念，概念是知识组织系统的基本单位；不同的概念组成领域词表，领域词表包含普通名词和专有名词类等不同种类的概念群；由领域词表与其标引和检索的文献共同组成知识系统。生物个体与概念实例对应，生物种群与概念对应，生物群落与领域词表对应，生态系统与知识系统对应。在结构和功能完全对应一致的基础上，将生态学的原理和方法引入到知识组织系统的功能和作用中，就形成了完整的知识组织生态系统。在 KOES 中，物种与概念对应，环境与文献对应，物种关系与概念关系对应，所有研究方向都是在这些基本的对应基础上开展和完成的。

2.2 实例个体的对应关系

个体层次的生态学，主要研究生物个体与环境之

* 本文系国家社会科学基金项目“面向叙词表构建的知识组织生态系统研究”(项目编号:15BTQ030)研究成果之一。
作者简介：常春(ORCID:0000 – 0003 – 2829 – 2589)，研究馆员，博士，E-mail: changchun@istic.ac.cn；杨婧(ORCID: 0000 – 0002 – 2254 – 7600)，助理馆员，硕士；李永泽(ORCID:0000 – 0002 – 1428 – 5735)，博士研究生。
收稿日期:2018 – 06 – 28 修回日期:2018 – 11 – 12 本文起止页码:146 – 150 本文责任编辑:易飞

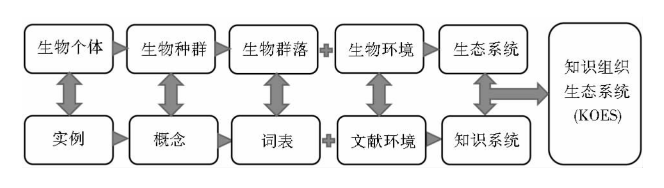


图 1　知识组织生态系统框架

间的相互关系，包括生物个体对环境的适应性，环境包括生物环境和物理环境。概念和实例是构成知识组织系统的重要成分，其中的实例（实体）与生态学中生物个体有着天然的对应关系，甚至对应的英文都是“individual”。例如草原上的一棵三叶草、一只羊、一头牛、一匹马、一只虎等，都是不同的生物个体，分别组成自己的物种种群，与生存的物理环境，例如温度、湿度、阳光和水等物理环境发生着联系。人名、地名、机构等实体名词，出现在文献标题、摘要、关键词、正文等不同位置，是代表概念的术语，每次出现都可以看作是一个实例个体。实例个体存在于文献环境中，不同实例个体组成各自的概念，与文献环境存在着词形、词义、语境、语义等关联关系。实例个体来源于文献环境，同时，也是文献环境的组成成分。

2.3　概念种群的对应关系

种群是栖息在同一地域中同种生物个体组成的集合。种群是由个体组成的群体，并在群体水平上出现了个体层次不具备的特征，例如出生率、死亡率、增长率、性别比、种内关系、种间关系、年龄结构、空间分布的等属性特征。例如草原上的羊群，群内有领头羊，头羊往哪个方向走，其他羊都会跟着它的步伐和方向走。羊群有种群的行为和特征，如以年为单位的小羊出生率和死亡率的统计、羊群年龄结构的统计、头羊的重要性等特征。在知识组织中，与种群对应的是概念，概念是思维的单元，是由实例个体组成的集合，概念可以是叙词表的优选词所代表的概念，也可以是分类法的一个类目或本体的一个类（class），概念同样有诞生、热点研究、停用特征，存在优选词与非优选词的等同关系，也存在着相关关系、等级关系等概念间关系，具有概念词频变化波动的概念成熟度判定，具有等同率、参照度、关联比等参数统计特征。物种种群与知识组织的概念有着完全一致的对应关系。

2.4　词表群落的对应关系

生物群落指特定空间内不同物种种群有规律的集合，是栖息在同一地域中的动物、植物和微生物组成的集合。在群落层次上，生态特征有群落的结构、演替、多样性、稳定性等。在陆地上有草原群落、湿地群落、森林群落、高山群落等。不同程度语义关系的 KOS，与生物群落相对应，词表同样有着宏观结构、微观结构、词表更新、概念多样性、概念稳定性等特征，有领域词表、综合词表的区分。领域词表与生物群落可以进行对应。

2.5　生态系统的对应关系

生态系统是在一定空间中生物群落和非生物环境的集合。主要研究内容集中在物质循环、能量流动、信息传递等方面。对应知识组织活动的系统性特征，在 KOS 的构建和应用中，同样存在知识传递、词表应用等系统性特征。地球上的能量流动表现在绿色植物（生产者）将光能转化为化学能，存储在有机物中；草食动物（初级消费者）通过对绿色植物的采食，将一部分能量转移到草食动物层；肉食动物（次级消费者）对草食动物进行捕食，使一部分能量又转移到肉食动物层。能量在生态系统中表现出单向传递、逐级递减的特征。人类在生产和社会活动中创造的知识，转化到出版社的著作、杂志社的期刊论文、会议的论文集等，经过网络渠道、图书馆等将知识传递给读者；知识组织体系的等级结构，从概念泛指到概念专指的属分关系传递过程，知识也有能量传递的特征。总之，KOS 的构建和应用过程，也符合生态系统的系统性特征。

3　知识组织生态系统的研究方向

3.1　知识组织生态系统研究方向

从实例个体、概念种群、词表群落和生态系统 4 个层次，分别介绍 KOES 的研究方向。详细内容如表 1 所示：

表 1　知识组织生态系统研究方向

领域	研究方向	领域	研究方向
实例个体	①文献环境实体名词挖掘研究 ②实例个体源于自然语言研究 ③实例个体词频统计变化研究 ……	词表群落	①词表概念可度量特征研究 ②词表概念多样性特征研究 ③词表概念稳定性特征研究 ……
概念种群	①个体概念成熟过程特征研究 ②同义近义概念种内关系研究 ③等级相关概念种间关系研究 ……	知识系统	①范畴等级结构的系统性研究 ②词族等级结构的系统性研究 ③知识传递能量流动特性研究 ……

3.2　实例个体层面研究方向

实例个体层面的研究方向限定在术语、语词等实体层面，主要研究语词的词形、词义规范化，例如优选词的遴选方法[6]；文献文本的自然语言处理，通过文献文本分词获取专业术语；通过数据挖掘自动获取人名、地名、机构名等专有名词实体数据；同义词[7]与同形异

义词的识别方法[8]。所有语词层面的研究都可以归入实例个体层面的研究方向。如果从知识组织体系应用和服务角度理解,个性化服务等用户个体相关的研究方向,也属于实例个体层面的研究方向。

3.3 概念种群层面研究方向

生物分类以物种种群为单位,KOS 的构建与应用以概念为单位,在概念层面,基于物种种群特征,可以开展大量研究,例如通过词频统计研究概念的成熟周期[9];通过概念的领域分布判断专业[10]概念和通用概念[11];同义近义概念种内关系研究[12];概念间等级关系的建立方法[13],概念间相关关系的建立方法[5]。以生物种群为单位,构成物种的视角下,对应概念为单位,进行 KOS 的构建与应用,研究概念的属性、概念的关系,存在大量有意义的研究方向。如果扩展到知识组织系统的应用角度,更可以包含用户本体构建、用户画像等对用户进行分类相关的研究方向。

3.4 词表群落层面研究方向

生物群落关注的是各类物种种群聚集成的集合,在多种生物组成群落的基础上,研究群落结构、生物多样性等特征。对应到 KOS,每部词表含有不同数量的概念,在词表层面,研究概念多样性问题;研究概念生态位空间结构,决定概念稳定性特征[3];研究概念的数量特征,例如参照度、关联比等各种语义关系计量参数等;研究概念关系组成的等级结构[4]或网络结构等;研究词表更新维护的概念动态演替等。所有不同概念群体间的关联和属性特征均属于词表群落层面的研究方向,如果扩展到知识组织系统服务领域,还会包括综合、专业、大学图书馆或文献机构的不同服务特征研究等。

3.5 生态系统层面研究方向

从系统性出发,研究生态系统物质循环、能量流动、信息传递等特征。对应到 KOS,同样存在方方面面的系统性特征,例如基于自然语言的叙词表自动构建研究[14];范畴体系和词族结构的系统性研究;词表系统之间的互操作问题研究[15];基于能量流动系统性的概念等级关系研究[4],等等。另外,生态系统层面除了系统性特征以外,也涉及到环境的因素,KOS 的基础应用就是对文献信息的标引和检索,因此,词表检索的相关研究属于生态系统层面的研究方向[16]。从知识组织系统服务角度出发,还会包括用户相关的各类系统性研究方向,例如用户获取知识的文化背景、信息基础设施、社会环境等。

4 知识组织生态系统研究相关讨论

4.1 知识组织生态系统学科属性

知识组织生态系统是应用生态学原理,从生态系统属性特征出发,将物种与概念对应、文献信息与生物环境对应,其中概念为知识组织系统的基本单位,物种为生物分类的基本单位,由实例个体、概念、词表、文献环境等因素构成相互作用、相互关联的知识组织生态系统。借鉴生态学方法,从理论上确立知识组织系统的构建和应用研究体系,解释知识组织系统的工作原理,为知识组织系统的构建和应用提供理论依据。KOES 概念属性属于知识管理范畴,研究内容包含知识相关的各类生命特征,但研究范围不包含具体的生命,研究对象没有真实的植物、动物和微生物,也不包含自然界的物质和能量,是将生态学的原理和方法应用到知识组织的过程,属于管理科学范畴。另外,鉴于知识形成、发展和应用的复杂性,知识组织生态系统是引进、借鉴生态学的原理和方法,并不是机械地对应和人为刻意地解释,随着研究的深入,知识组织生态系统在一些方面会体现出其更复杂的特点。

4.2 种群个体与概念实例的对应关系

对于生物种群,一个物种可以形成一个种群,一个种群通常含有多个个体,个体规模大小直接关系到该物种在生态系统中的作用和地位。例如草原上的羊群,羊是一个物种,羊群中每只羊是一个个体,羊群规模大小,对羊的繁盛、对草原生态系统的作用,都有重大影响。物种与知识组织体系的概念相对应,概念可以使用规范化的术语作为标签,代表一个概念,概念同样应该有多个实例,那么,在知识组织生态系统中,如何体现概念的实例呢?通过研究探索,我们的结论是概念优选词在文献中每出现一次,本次词频就等同于概念的一个个体或一次实例,例如一篇论文,研究主题是“本体”,本体在标题、摘要、关键词 3 个位置都出现了,正文中也出现了 20 次,这样,在这篇论文中,“本体”是一个概念,“本体”在这篇论文中有 23 个概念实例或有 23 个个体,论文中每个“本体”出现的位置不同,代表的属性也是有区别的,例如在标题、摘要、关键词、结论等位置的“本体”,对论文的主题表达贡献要大一些,而在综述、展望等部位的“本体”,对主题的贡献要小一些。如果是对于含有大量文献的数据库,我们把每篇主题是“本体”的文献当作是一个概念个体,例如,在 CNKI 网络数据库中,含有主题词“本体”的文献为 47 486 篇[17],可以认为,“本体”概念有 47 486 个

常春，杨婧，李永泽. 知识组织生态系统构架形成与研究进展[J]. 图书情报工作,2019,63(7):146－150.

个体,例如《基于 Logistic 种群增长规律的概念词频变化研究》一文就是这样处理的[9]。

4.3 物种关系与概念关系的对应

动物－哺乳动物－牛是概念包含关系,生物中的概念包含关系与知识组织系统中的概念等级关系是完全对应的,更具体的是属种关系,“哺乳动物”一部分是“牛”,“牛”全部是“哺乳动物”,是典型的属种关系。生物物种之间存在着捕食、竞争、寄生和共生关系,通过方向性和依赖性可以进行区分,捕食是从捕食者到被捕食者的单向过程,但食物有替代性,不一定只依靠一种食物;竞争是相互的,不会是单一方向的,相互之间也没有依赖关系;寄生是寄生物寄生到寄主的体表或体内,是单向依赖寄主生存的;共生双向依赖的,是对双方生物都有利的。以上生物种间关系也可以与相关关系的种类进行对应,我们将叙词表国际标准 ISO 25964－1 的相关关系不同类型进行分析和测试,均可以唯一地归类到一种种间关系中。同时也使用了《汉语主题词表(工程技术卷)》的相关关系进行了实证,结论是所有的相关关系都可以唯一地归类分为捕食、竞争、寄生和共生类型中的一种[5]。例如,单向依赖的概念关系可以对应物种间的寄生关系,“电流”与“安培”是相关关系,“电流”是客观存在的,先有“电流”,“安培”是“电流”的计量单位,“安培”是单向依靠“电流”而生,没有“电流”,其单位“安培”也不会存在,因此是单向依赖型的寄生关系。在知识组织系统中,如果描述一个概念与另一个概念是寄生关系,感觉难以理解,有无通用于知识组织的术语描述这种关系,替代生物种间的捕食、竞争、寄生和共生关系,也是一个需要探索的研究方向。

4.4 研究单位的聚焦和相对性

物种是生物分类的基本单位,例如马、牛、羊分别为不同的物种;概念是知识组织系统的基本单位,例如概念是叙词表的基本构成单位。在知识组织生态系统的不同层面,基本单位与研究单位可以限定在不同层面,例如,如果限定在概念个体层面,研究单位就是实例个体,例如优选词的遴选方法[6],就是针对术语层面,研究术语的规范化问题,研究同义词[7]与同形异义词的识别方法[8]。如果研究单位是概念层面(物种),则研究概念的属性,例如概念成熟周期[9];如果研究单位是词表层面(群落),则研究概念生态位空间结构,决定词表稳定性特征[3];研究词表的数量特征,例如参照度、关联比等各种语义关系计量参数等;如果研究单位是知识系统(生态系统),则研究内容包括词表系统之间的互操作问题[15]、词表检索相关研究等[16]。可见,研究单位大小范围限定不同,研究的问题也是不一样的。

5 结语

生态学作为一门独立的一级学科,经过科学总结,具有一般的原理和方法,这些原理和方法已被广泛应用于社会科学各个领域。叙词表、分类法、本体等知识组织体系,在其构建和应用中,需要进行概念遴选,进行概念关系的建立,可以对文献信息进行标引和检索。所有这些活动与生态学有着完整一致的对应关系,知识组织体系的概念实例与生物个体相对应,概念与生物种群对应,领域词表与生物群落对应,文献环境与生物环境对应,知识系统与生态系统对应。在结构和功能上高度一致对应基础上,本文将生态学的原理与方法,引入到知识组织系统的构建和应用中,分别在概念实例层面、概念层面、领域词表层面和知识系统层面开展了研究,其中,概念实例层面是从语词的词形规范、词义规范方面展开;概念层面是研究概念内部的同一种群内部个体之间的关系,也研究概念之间物种间关系;词表层面是研究类似生物群落一样概念集合的特征属性;知识系统层面主要从系统性出发,研究词表的检索功能等应用等内容。

参考文献:

[1] 牛翠娟,娄安如,孙儒泳,等. 基础生态学[M]. 3 版. 北京:高等教育出版社,2015.
[2] 常春. 面向叙词表构建的知识组织生态系统研究[J]. 图书情报工作, 2016, 60(15): 101－107.
[3] 杨婧, 常春. 基于生态位法则的概念稳定性研究[J]. 图书情报工作, 2016, 60(13): 27－32.
[4] 李永泽, 常春. 基于生态学能量传递的词族层次结构研究[J]. 情报杂志, 2017, 36(3): 161－165,172.
[5] 李永泽, 常春. 基于生态学种间关系的叙词表相关关系分类研究[J]. 图书情报工作, 2018, 62(8):123－129.
[6] 陈白雪,常春,刘春燕,等. 叙词表概念优选词选择方法研究[J]. 情报杂志, 2015, 34(12): 170－175.
[7] 王刘安, 常春. 用代传导中同义术语识别研究[J]. 情报理论与实践, 2014, 37(9): 97－100,91.
[8] 陈白雪,常春. 同形异义词机器辅助识别方法研究[J]. 数字图书馆论坛, 2015(5): 8－13.
[9] 杨婧, 常春. 基于 Logistic 种群增长规律的概念词频变化研究[J]. 情报科学,2017,35(8): 15－18.
[10] 常春, 赖院根. 专业概念机器辅助分类方法研究[J]. 现代图书情报技术, 2011(10):34－39.
[11] 常春, 赖院根. 数字环境下通用概念获取方法[J]. 图书情报

工作, 2011, 55(23): 22 –25.
[12] 张冰, 常春. 基于术语原形化的英文同义词群构建方法研究[J]. 情报杂志, 2014, 33(7): 171 –175.
[13] 邓盼盼,常春. 叙词表编制中的概念多重属分关系研究[J]. 情报科学,2015,33(5):59 –62.
[14] 杜慧平,何琳,侯汉清. 基于聚类分析的自然语言叙词表的自动构建[J]. 国家图书馆学刊, 2007(3):44 –49.
[15] 常春, 曾建勋, 吴雯娜, 等.《汉语主题词表》与英文超级科技词表概念映射构架设计[J]. 数字图书馆论坛, 2012 (12): 28 –32.
[16] 常春. 叙词表的术语服务方式研究[J]. 图书情报工作, 2012, 56(22): 12 –15.
[17] 本体[EB/OL]. [2018 –06 –10]. http://kns.cnki.net/kns/brief/result.aspx? dbprefix = SCDB.

作者贡献说明:
常春:负责论文构思和撰写;
杨婧:负责数据采集与实证;
李永泽:负责数据采集与实证。

Formation and Research Progress of Knowledge Organization Ecosystem

Chang Chun[1] Yang Jing[2] Li Yongze[3]
[1] Institute of Scientific and Technical Information of China, Beijing 100038
[2] Chinese Version Library, Beijing 100005
[3] School of Information Management, Nanjing University, Nanjing 210023

Abstract: [**Purpose/significance**] By introducing the principles and methods of ecology into the study of knowledge organization system, the functions and actions of knowledge organization system can be brought into full play. [**Method/process**] Based on the research results of ecology, the principles and methods of ecology are introduced into the construction and application of the knowledge organization system. The biological individual corresponds to the concept individual, the species population and the concept correspond, the biological community and the vocabulary correspond, the biological environment corresponds to the literature environment and on the basis of the complete and consistent correspondence of the structure and function, the working principle and method of the knowledge organization system are studied. [**Result/conclusion**] The framework of the knowledge organization ecosystem is refined and perfected, and the main contents of each level are analyzed and summarized in the four aspects, including the concept instance, the concept population, the concept community (vocabulary) and the ecosystem.

Keywords: knowledge organization ecosystem　ecosystem　knowledge organization system　thesaurus